EXHIBIT 2

THE COST OF THE LOCAL TELECOMMUNICATION NETWORK

A Comparison of Minimum Spanning Trees and the HAI Model

by

Christian M. Dippon
National Economic Research Associates

and

Kenneth E. Train
Department of Economics
University of California, Berkeley

August 26, 1998

Abstract: Under the Telecommunications Act, estimates of local distribution costs may be used to help quantify the subsidy for specified local services whose costs exceed their tariffed rates and as a guide for the pricing of unbundled network elements. The most widely-circulated model for estimating these costs, the HAI model, uses a particular procedure to calculate the distribution network and cable length that is required to serve a cluster of customers. We compare the HAI procedure with the minimum spanning tree (MST), which gives the shortest distance for connecting a set of locations. For each cluster in Minnesota we calculated the distribution length with the HAI procedure and the length of the MST. We find that the HAI length is shorter than the MST length in 77% of the main clusters. In low-density areas, the HAI length is less than the MST length for 81% of the main clusters.

THE COST OF THE LOCAL TELECOMMUNICATION NETWORK

I. INTRODUCTION

The implementation of the Telecommunications Act of 1996 entails measurement of the costs of supplying telecommunications services for at least two distinct purposes: (i) pricing unbundled network elements and interconnection and (ii) quantifying the subsidy contained in current prices for local exchange services. The Federal Communications Commission (FCC, 1997) has specified that these costs are to represent forward-looking efficient costs – that is, the costs that would be incurred by a carrier that provided these elements and services using least-cost methods at today's prices and technology. Since these costs are not directly observable, "cost proxy models" have been developed to estimate them. The cost proxy model that has been submitted most frequently in regulatory proceedings is the HAI model (1998).

Much of the debate on the adequacy of such models for the purpose for which they are offered is centered on how the models determine the local distribution network.[1] The current network may not be the most efficient for serving today's known demand since, e.g., the network was constructed incrementally over time as demand grew rather than being configured optimally for the current number and location of customers. Therefore, instead of costing the current distribution network, the models construct a hypothetical network that links each customer to a wire center. The location of customers and the layout of the network that connects customers to the wire centers have a significant impact on the estimated costs, because these assumptions determine the predicted loop length.

The HAI model uses a particular procedure, described below, for determining the distribution network and thus the total cable required to serve customers. Recently, several parties have suggested that this procedure under-estimates the amount of cable that is needed to link customers to the wire centers, leading to lower-than-accurate costs (Prisbrey, 1998;

[1] The local distribution network is a component of the local exchange network. It addresses the portion of the network extending from Serving Area Interfaces (SAIs. i.e., the interconnect between feeder and distribution) to the customers' premises.

Staihr, 1998). We investigate this issue empirically by comparing the HAI distribution network procedure to the minimum spanning tree (MST). A MST is a mathematical graph theory construct used to connect a set of points at the least possible length of total connecting lines (see, e.g., Flood, 1977; Biggs, 1994.) As such, the MST provides a lower limit (subject to caveats described below) to the distribution cable that is needed to serve a cluster of customers.

Our analysis is performed for GTE's territory in Minnesota. The HAI model identifies over a thousand clusters of customers in the territory. For each cluster, we calculate the total cable length implied by the minimum spanning tree and compare this length with that obtained by the HAI procedure. We find that lengths in the HAI model are considerably less than the length of the minimum spanning tree. This result implies that the HAI procedure provides less distribution cable, and hence lower costs, than is physically possible to use in serving the customers.

Our findings are consistent with, and generalize, those of Staihr (1998). He reported that, for several clusters of customers in Nevada, the distribution cable obtained by the HAI procedure is less than that of the minimum spanning tree, by as much as a factor of nine. Mercer and Klick (1998) responded to his finding by stating that his analysis examined only a very few clusters in one area such that the results cannot be considered representative. Our analysis examines a different area and includes all clusters in the area.

II. DISTRIBUTION NETWORK IN THE HAI MODEL

The HAI model identifies the locations of customers and then groups these locations, following a set of engineering rules, into clusters. In our comparison with minimum spanning trees, we take the location and clustering of customers as given by the HAI model, and therefore do not describe here the location and clustering process. Given a cluster of locations, the HAI model determines distribution network in a series of steps. We describe these steps through an example. Figure 1 illustrates a set of customer locations that constitute a cluster.

Figure 1: Locations in a cluster

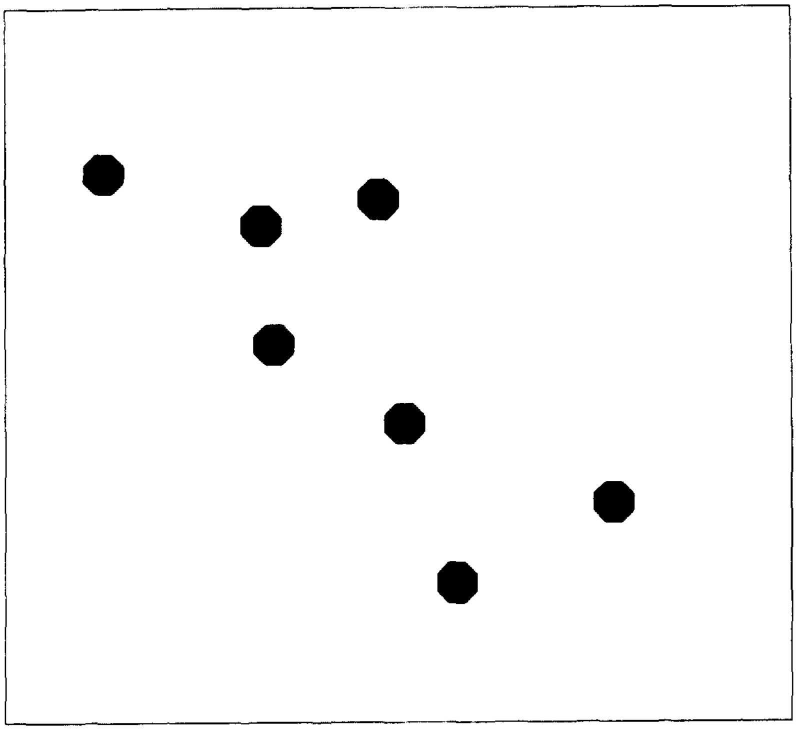
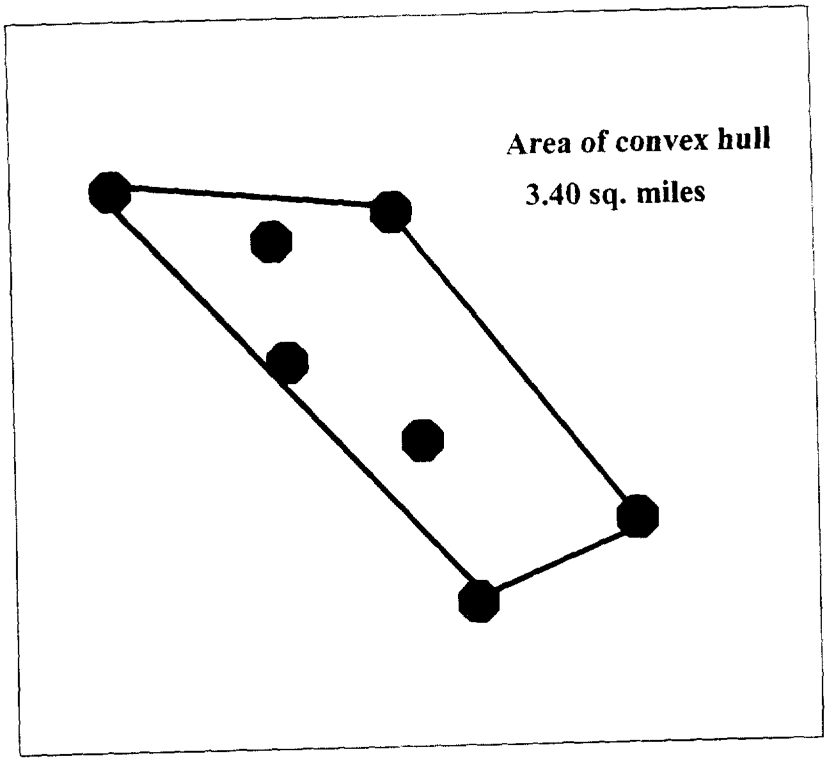


Figure 2: Convex Hull around locations

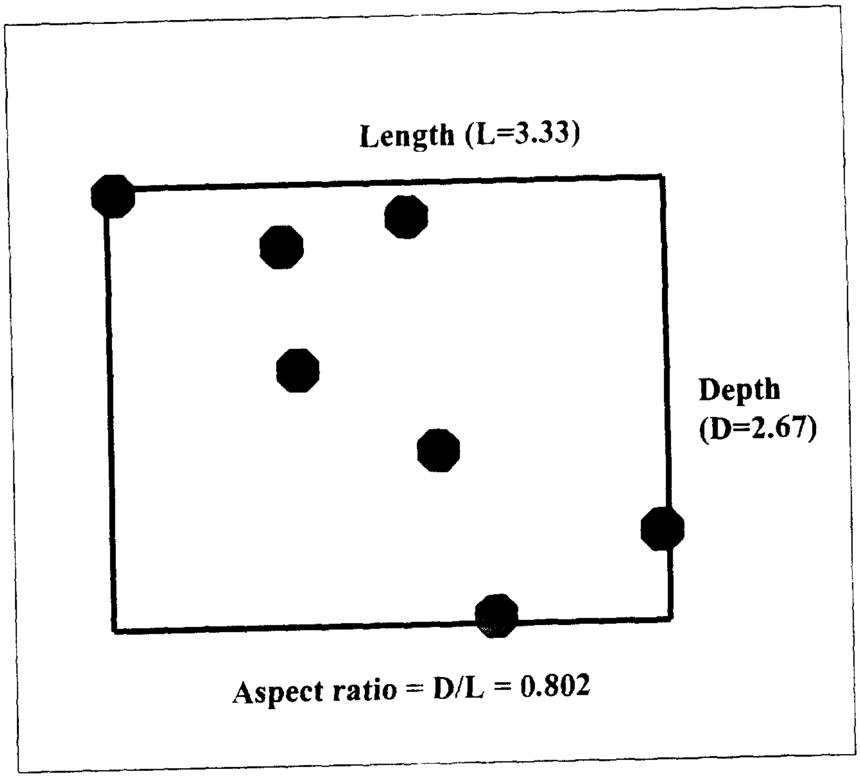


The first step of the HAI procedure is to determine the convex hull[2] that contains all of the locations in the cluster. This convex hull is shown in Figure 2. The area of the convex hull is calculated; in our example, the area of the convex hull is approximately 3.40 square miles.

Next, the process determines the smallest rectangle that contains the locations, called the "minimum bounding rectangle" (MBR). This MBR is shown in Figure 3. The "aspect ratio" of the MBR is calculated by dividing the length of the MBR by its depth. In our case the aspect ratio is 0.802.

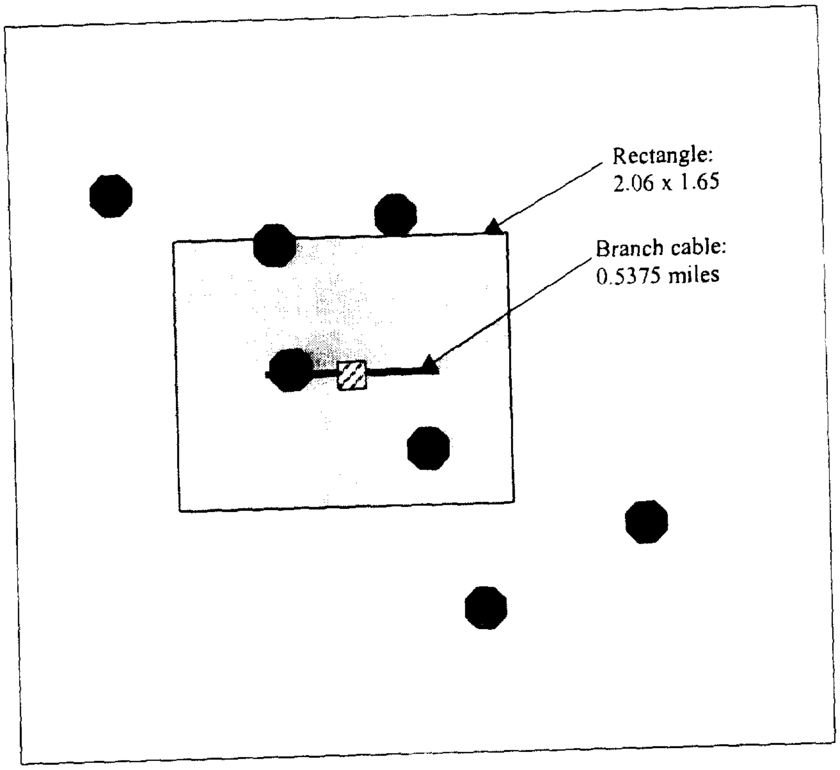
[2] The convex hull of a set is the smallest convex set containing it. A set is convex if, given two points in the set, the straight line segment joining the two points is also contained in the set.

Figure 3: Minimum Fitting Rectangle around locations



The algorithm then constructs a rectangle with the same area as the convex hull (3.40 sq. miles) and the same aspect ratio as the minimum bounding rectangle (0.802). This rectangle for our example, depicted in Figure 4, is 2.06 miles wide and 1.65 miles deep.

Figure 4: Distribution for Cluster



Intuitively, the distribution network is designed as if the customers were located on equal-sized lots within this rectangle, with each lot being twice as deep as wide. In our illustration, each lot is 0.493 miles wide and 0.986 miles deep. The lots cannot be shown on Figure 4 because they do not fit into the rectangle. Lots can only be represented within the rectangle when the number of locations and the aspect ratio of the rectangle allow such representation. For example, if there were eight locations in the cluster and the aspect ratio for the rectangle were 1.0 (half the aspect ratio of the lots), then the rectangle would be filled by two rows of four lots each. Similarly, if there were 16 locations and the aspect ratio were 2.0, the rectangle would be filled with four rows of four lots each. Generally, however, the aspect ratio and number of locations do not enable the lots to be represented as rows in the rectangle.

The distribution cables consist of backbone cable, branch cable, and drop cable. The length of backbone cable is calculated as the depth of the rectangle (1.65 in our illustration) minus the depth of two lots (2 times 0.986), with a minimum of zero. With several rows of lots within the rectangle, this length corresponds to having backbone cable run vertically to connect the top and bottom rows (i.e., from the bottom edge of the top row of lots to the top edge of the bottom row of lots.) The same formula is applied when, as in our illustration, the lots do not fall in rows that fill the rectangle. In our illustration (as when there are two or fewer rows of lots), there is no backbone cable.

The number of branch cables is calculated as the depth of the rectangle divided by the depth of two lots, rounded up to the next higher integer, times two. With lots falling into rows that fill the rectangle, this number corresponds to having branch cables running horizontally between every two rows of lots, with one cable running "east" from the middle of the rows and another cable running "west" from the middle. As with backbone cable, the formula applies for any number of lots and aspect ratio. In our illustration, there are two branch cables.

The length of each branch cable is half of the width of the rectangle (in our illustration, 2.06 divided by 2) minus the width of a lot (0.493). With lots in rows, this length corresponds to the distance from the middle of the rows to within one lot of the end of the row. In our illustration, each branch cable is 0.5375 miles long. The two branch cables have a total length of 1.075.

Finally, drop cables connect each customer to the branch cables. The length of the drop cables is determined within the HAI procedure through look-up tables based on the population density of the census block that contains the cluster. The density zone of the cluster is not used at this point. For our illustration, we assume that the cluster falls in a low-density category, for which the drop lengths are 150 feet. With seven customer locations, the total amount of drop cable is 0.20 miles (150 feet x 7 / 5280 feet per mile.)

The total distribution length resulting from this process is the sum of the length of the backbone cable, the branch cables, and the drop wire for all seven locations:

Total Backbone Cable Length =	0
Total Branch Cable Length =	1.075 miles
Total Drop Cable Length =	0.20 miles

Total Distribution Length = 1.275 miles

III. MINIMUM SPANNING TREES

Minimum spanning trees have also been called shortest path or minimal connector problems (e.g., Chartrand, 1985.) The issue is common for railroads where the goal is to construct a railroad system at least cost. Other common applications include computer circuits, long-distance telephone lines, delivery routes and mail routings that seek to find a minimum total length of routes that will connect all desired locations.

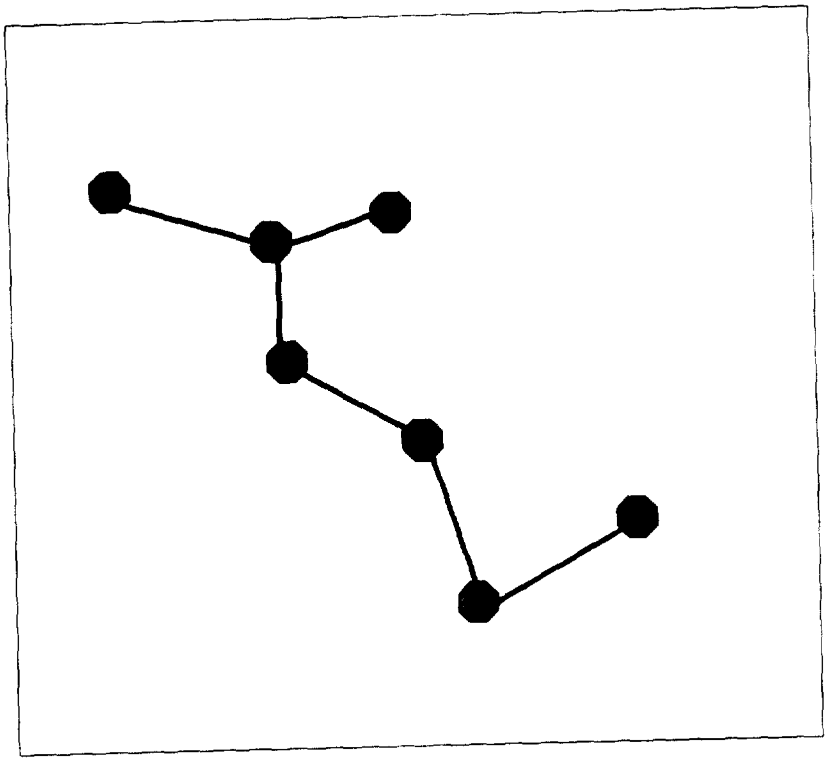
Numerous algorithms have been developed to compute a MST. Perhaps the simplest algorithm, at least from a pedagogical perspective, is Kruskal's (see, e.g., Sedgewick, 1988). Start by calculating the distance between each pair of points and rank-ordering the pairs in terms of the distance between them, from closest to farthest. Identify the first pair of points on this list (i.e., the pair that are closest) and connect them with a line. These two points constitute the current spanning tree and the line between them is called an "edge." Now consider the second pair on the list and connect them. If one of the points in the second pair was also in the first pair (e.g., the first pair is <a,b> and the second is <b,c>), then the current spanning tree has three points in total (a,b,and c), with two edges connecting the three points (from a to b and from b to c.) Otherwise (i.e., if the points in the second pair are distinct from those in the first pair, such as <a,b> and <c,d>), then the current spanning tree is disjoint, consisting of four points (a,b,c,d) with two edges connecting the points in each pair (from a to b and from c to d) but without the two pairs being connected to each other (a and b are not connected to c and d.) Consider now the third pair on the list. Both points in this pair might already be in the spanning tree (for example, if the third-ranked pair is <a,c> in either of the cases above.) In this case, do not extend the spanning tree and move onto the fourth-ranked pair. If, on the other hand, at least one of the points is not in the current spanning tree (e.g., if <d,e> is the third-ranked pair), then connect the pair with a line, such that they become part of the current spanning tree. Move to the fourth-ranked pair, and do the same: if both points are already in the current spanning tree, ignore the pair and move on; otherwise, connect them, thereby expanding the spanning tree. Continue until all points have been added to the spanning tree. Note that, while all points in the spanning tree will be connected at the end of the algorithm, in intermediate steps the Kruskal algorithm can be working on many disjoint sections of the tree. These sections are joined before the algorithm completes. This algorithm is a type of a "greedy" algorithm, because it chooses at each step the shortest edge to add to the MST.[3]

The algorithm developed by Prim (1957) builds a tree by adding one point at a time to the current spanning tree. Start by connecting the two closest points. Of the remaining points, find the one that is closest to either of the two points in the current tree. Connect this third point to whichever of the first or second points it is closest to, thereby expanding the tree to three points. Continue this procedure, finding the shortest distance between any already-connected point and any not-yet-connected point, and connecting those points. Note that the current spanning tree is fully connected in each step, unlike the Kruskal procedure which can have disjoint sections in intermediate steps. While the MST resulting from different methods might differ, there is only one shortest total length. That is, the length of the MST will be identical regardless of algorithm.

Figure 5, illustrates the minimum spanning tree for our example of customer locations.

[3] Greedy algorithms make locally optimal choices, choosing the best alternative at each step in the process (Biggs, 1994). In many circumstances, MST's being an important case, a sequence of locally optimal choices results in a globally optimal solution.

Figure 5: Minimum Spanning Tree



There are two limitations to the MST in context of determining the forward-looking costs of a telephone network. First, telephone networks are not constructed by directly linking one subscriber. Rather, cable is run along a path, and short drops from terminals connect this cable to subscribers. Those terminals represent additional points in the network, and if optimally placed, the addition of these points can reduce the total length of the distribution network. Stated in terms of spanning trees, the MST for a given set of points may have greater total length than the MST for these points augmented by some appropriately chosen other points. For example, consider four points at the corners of a 1x1 square. The MST for these four points traces out three sides of the square, for a total length of 3. However, the four points can be also connected with two diagonal lines from each corner to the opposite corner, using a connection point in the middle of the square (where the two diagonals intersect.) The length of each diagonal is SQRT(2), such that the total length is 2*SQRT(2), which is less than the length of the MST without adding points.

The construct that calculates a MST by allowing extra nodes is known as a Steiner Minimum Tree (Cournant, 1941; Balakrishnan, 1989). Du and Hwang (1990) proved a 22-year-old conjecture of Gilbert and Pollak (1968) which says that the minimum ratio between the length of a Steiner minimum tree and a minimum spanning tree is SQRT 3/2. That is, adding extra interconnection points cannot reduce the total length of the tree by more than about 13 percent. In most situations, the difference is much less than 13 percent. For the example of the 1x1 square given in the previous paragraph, the Steiner tree is six percent shorter than the MST without the additional point.

The second limitation of MSTs for representing telecommunication networks is that the line segments of a MST run directly from one point to another. They do not account for geographical obstacles such as rivers, mountains, lakes, freeways, rights-of-way, etc. The MST is calculated using airline miles (the straightest distance between points) rather than the actual amount of cable (route miles) that would be required to connect customers given the actual geography. This issue arises in many contexts regarding telecommunication networks, and commonly used air-to-route ratios conversion factors have been developed. Generally, airline miles for cable are converted to route miles using ratios between 1.3 and 1.6, with 1.414 being perhaps the most common.[4] That is, the minimum possible length of cable to serve a group of customers is generally 30% or more greater, due to geographical obstacles, than the length of the MST.

IV. COMPARISON

The definition of clusters and the identification of customers' locations within clusters are specified by the HAI model, and we take them as given. We examined all clusters in GTE's territory in Minnesota. For each cluster, we identified the MST using the Prim algorithm[5] and calculated the total length of the tree. We compare this number to the total

[4] The route-to-air ratio of 1.414 is simply the square root of 2. It assumes that instead of direct connection of two points, the points are joined rectilinearly.

[5] We used a program developed by Stopwatch Maps, Inc., a Geographic Information System consulting firm in St. Louis, Missouri.

distribution cable length, including drop wires, that is produced by the HAI model.[6] Our results are given in Table 1 and summarized below.

Table 1: Distribution of Ratio of HAI Length to MST Length				
	Total number of clusters	Number of clusters with ratio below 1	Number of clusters with ratio below .87	Number of clusters with ratio below .5
Main clusters	890	682	620	258
Other (outlier) clusters	248	114[7]	114	114
Main clusters in lowest two density zones[8]	839	680	618	258

The HAI model distinguishes "main" clusters, which contain five or more locations, and other clusters which contain four or fewer locations.[9] For the main clusters, the total distribution length from the HAI procedure is less than the length of the MST in 682 (77%) of the 890 clusters in Minnesota. For 620 (70%) main clusters, the HAI length falls below the

[6] We calculated the drop length in the HAI model using two different methodologies: 1) by adding to the HAI reported total distribution length the total drop length based on the *number of locations*; and 2) by adding to the HAI reported total distribution length the total drop length based on the *sum of total number of households and businesses*. The results for the two methods of calculating drop cable length differed only minimally (often less than 1 percent); thus, we report our findings based on the first methodology only.

[7] For outliers with one location the length of the MST is by definition zero. In these cases, HAI's distribution length is necessarily longer.

[8] The lowest two density zones in HAI consist of areas that contain 0-5 lines per square mile and 5-100 lines per square miles.

[9] Main clusters account for 74 percent of the customer locations and 80 percent of the distribution cable in the HAI model.

MST length by more than 13 percent; that is, for 70% of the main clusters, the HAI procedure gives shorter lengths than could be physically possible even if one assumes that there are no physical impediments, such that the air-to-route ratio is one, and that the theoretically maximum benefit of adding connection points for a Steiner tree is attained. In some clusters, the HAI procedure produces length estimates that are less than 10% percent of the MST length.

For other (non-main) clusters, 117 (47%) of the 248 clusters have lengths from the HAI procedure that are shorter than the MST length. For all of these 117 clusters, the HAI length is less than half of the MST length.

The FCC and several state commissions are attempting to use cost proxy models to assist in determining the amount and allocation of the Universal Service Fund (USF.) Clusters in low-density areas, where costs are generally above tariffed rates, are particularly important for determining the subsidy under USF. Considering all the main clusters in the two lowest density categories, the HAI procedure gives a total distribution length that is less than the MST length in 680 (81%) of the 839 clusters. The HAI length is more than 13% less than the MST length in 618 (74%) of these clusters.

As a concluding word, it is important to note that the distribution cable lengths in the HAI model, as in all cost proxy models, have an impact on many of outputs beyond the investment cost of the distribution cable itself. Support structures (e.g., poles, manholes, trenches, conduits, pull boxes), maintenance costs, associated power and back up power equipment, and many other factors are affected by the layout and length of the distribution network. Underestimation of the network length can lead to seriously insufficient support for explicitly subsidized services and incorrect prices (and hence incorrect price signals for investment and entry) for unbundled network elements.

References

Balakrishnan, V.K., 1997, *Schaum's Outline of Graph Theory*, McGraw Hill Publishing Company, New York.

Biggs, N.L., 1994, *Discrete Mathematics*, Clarendon Press, Oxford.

Chartrand, G., 1985, *Introductory Graph Theory*, Chapter 4, Dover Science Books, Dover Publications, Inc., New York.

Courant, R.& Robbins, H., 1941, *What is Mathematics?*, Oxford Univ. Press, New York.

Du, D.Z. & F. Hwang, 1990, "A Proof of Gilbert-Pollak's Conjecture on the Steiner Ratio", Publication 90-72, Center for Discrete Mathematics & Theoretical Computer Science of Rutgers University.

Federal Communications Commission, 1997, "Report and Order, In the Matter of Federal-State Joint Board on Universal Service," CC-Docket No. 96-45, Adopted: May 7, 1996, Released: May 8, 1997, paragraph 233.

Flood, J.E., ed., 1977, *Telecommunication Networks*, IEE Telecommunication Series 1, Section 13.4.

Gilbert, E.N.& Pollak, H.O., 1968, "Steiner Minimal Trees", *SIAM J. Appl. Math*, Vol. 16, No. 1, 1-29.

HAI, Inc, 1998, *HAI Model 5.0a, Model Description*, Boulder, Colorado, February 28.

Mercer, R.A., and J.C. Klick, 1998, "Supplementary Reply Testimony, Before the Public Utility Commission of Texas," PUC Docket No. 18515, June 10.

Prim, R. C., 1957, "Shortest Connection Matrix Network and Some Generalizations," *Bell System Technical Journal*, Vol. 36, 1389-1401, November.

Prisbrey, J., 1998, "A Test of Customer Dispersion in the HAI Customer Location Algorithm," Memorandum, Federal Communication Commission, May 13.

Sedgewick, R., 1989, *Algorithms*, Addison-Wesley Publishing Company, Inc, New York, reprinted with corrections, August.

Staihr, B., 1998, "Ex-Parte Submission," Federal-State Joint Board on Universal Service, CC Docket No. 96-45, April 20.